WUYI CITATION TONE ACOUSTICS: PROBLEMS FOR TONOLOGICAL REPRESENTATION

Napier Guy Ian Thompson

Phonetics Laboratory, Department of Linguistics (Arts),
Australian National University

ABSTRACT This paper describes the acoustic characteristics (F_0 and duration) of the citation tones of Wuyi, a Southern Wu dialect, belonging to the Wuzhou dialect sub-group. Mean acoustic data from one male speaker are presented. It is shown how the results of the analysis pose questions for tonological representation in current tonological theory.

INTRODUCTION

The Wu dialects are spoken chiefly in China's Zhejiang Province, located on the central eastern coast, and belong to the Sinitic branch of Sino-Tibetan. Of particular interest among the Wu dialects is the dialect sub-group known as Wuzhou, which belongs to the Southern Wu dialects. The Wuzhou dialect sub-group forms a region situated fairly centrally within the Zhejiang province. The dialects spoken within the Wuzhou region are interesting for a number of reasons. Firstly, Wuzhou dialects are renowned for having a high number of citation contrasts (Zhu & Rose, 1998) and Wu dialects, in general, are renowned for what is possibly some of the most complex tone sandhi in the world (Rose, 1990). Secondly, the region is very interesting in that there is such a high degree of heterogeneity, especially with regard to tonetics and tonology, exhibited by dialects of relative geographical proximity. The dialect we are concerned with in this paper is that of Wuyi. Wuyi is situated in the central southern region of the area bounded by the borders of Wuzhou. Geographically, it is located just south-west of the Wuyi River, lying approximately 300 km south-west by south of Shanghai and about 30 km southeast by east of the major city Jinhua, which is located in the central west of the Wuzhou region.

Preliminary auditory analysis of data on the Wuyi dialect indicates that it has what appear to be three rising tones in contrastive distribution on citation syllables. If this is indeed true, it poses serious problems for representation, as it is not possible to keep three rising tones separate under the usual phonological framework, such as Yip (1980). The purpose of this paper is to confirm or refute this observation by presenting a description of the Wuyi citation tone system in terms of mean acoustic data.

Traditionally, Chinese linguistics recognises 8 historical tonal categories, which represent assumed contrasting tonal contours present in late Middle Chinese. The eight categories are 'yinping' or (1a), 'yinshang' (2a), 'yinqu' (3a), 'yinru' (4a), 'yangping' (1b), 'yangshang' (2b), 'yangqu' (3b) and 'yangru' (4b). There exists also a distinction between stopped and unstopped tones, whereby stopped tones are referred to as 'ru' and unstopped tones are collectively referred to as 'shu'. It is useful to refer to these historical categories when discussing both synchronic and diachronic tonal differences within many Chinese dialects and this system of reference will be used throughout this paper. This system is interesting in itself in that it seems to embody the modern tonological concepts of Yip's (1980) Tonal Registers and also to a degree, the concept of tone pairing.

Wuyi appears to contrast 8 distinct tone shapes on citation syllables. The following is a description of the auditory characteristics of the Wuyi citation tones: **Tone 1a** has low rising pitch [14]. **Tone 1b** has mid fall-rising pitch, falling approximately from the centre of the speaker's pitch range [313]. **Tone 2a** has high level-rising pitch [334]. **Tone 2b** has mid level-rising pitch [223]. **Tone 3a** has high falling pitch [51]. **Tone 3b** has mid rise-falling pitch [231]. **Tone 4a** has truncated high falling pitch [$53^{ʔ}$]. **Tone 4b** has truncated mid fall-rising pitch [$312^{ʔ}$]. Thus, the Wuyi dialect exhibits 3 rising tones, 2 falling tones, 2 concave tones and 1 convex tone.

PROCEDURE

The corpus used in this analysis consists of 40 monosyllabic words (5 words belonging to each of the 8 historical tone categories). The sole restriction on syllable structure was that imposed by the language, thus, syllables may have the structure C(G)V({N/ʔ}) where C can be any permissible consonant including glides and [ʔ], V can be any permissible vowel or diphthong including buzzed vowels and the permitted appearance of (G) in the second onset slot is dependant upon the quality of C. The full corpus is presented in Thompson (in preparation). The informant, a male speaker aged 23 years at the time of recording, produced three tokens for each of the 40 monosyllables with pausing between each token, totaling 120 utterances. The analysis ignores the final token in each group to control for any list-final intonation. Categorical shift also prevented the use of some words in the categories for which they had been recorded. Thus, each contour in the analysis represents the mean of between 4 to 6 words × 2 tokens each = 8 to 12 utterances, except the ru contours 4a S and 4a P (see section 'Wuyi Stopped ('Ru') Citation Tones') which are the mean of 4 utterances and 6 utterances respectively.

Aligned audio waves and wide-band spectrograms were made to facilitate segmentation of the data. F_0 samples were taken at 10% points across a time base which was determined as the distance from the offset of the first syllabic onset slot to the offset of phonation, following the findings of Rose (1998) that sonorant consonants in secondary position carry F_0 relevant to Tone, whereas those in initial position do not. Dummy consonants, whose duration was determined by taking mean durations over a number of segmentable sonorant initials, were used on non-segmentable material such as glide initials and such structures as shown in Figure 1. Duration measurements of this interval were also taken. Arithmetical mean and standard deviation values were calculated for all contours.

CV
V
ŋ

Figure 1.

RESULTS

Table 1 presents the mean Fundamental Frequency (F_0) of the Wuyi citation contours at 10% sample points across total tonal duration. Total tonal duration is also presented in the right-most column. All F_0 values are in Hz to the nearest Hz and all duration values are in centiseconds to the nearest 10th.

TONE	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	DURATION
1a	122	118	115	114	115	117	123	133	143	152	160	25.0
1b	148	143	136	129	123	120	120	124	131	139	148	28.9
2a	152	154	155	155	155	157	159	161	163	166	168	22.2
2b	134	135	135	135	135	136	139	143	146	147	143	25.1
3a	177	178	178	174	166	155	145	138	133	127	126	16.1
3b	133	136	142	146	147	146	142	135	130	126	124	21.5
4a S	174	173	171	170	169	167	166	164	161	159	156	10.3
4a P	184	181	177	174	170	166	158	151	141	135	130	14.8
4b	143	136	132	126	123	122	124	128	133	133	131	25.0

Table 1. Wuyi Citation Tones: Mean F_0 Contours

TONE	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	DURATION
1a	5.7	5.1	5.4	4.4	4.0	3.9	7.0	9.2	8.4	5.7	6.3	1.3
1b	6.6	7.6	7.9	8.0	5.9	5.4	5.5	5.9	7.4	6.3	9.8	1.9
2a	5.8	5.3	5.6	5.5	6.1	7.2	7.8	6.8	5.9	4.8	5.1	1.9
2b	4.6	5.4	6.0	6.9	7.1	7.4	6.7	7.0	7.8	10.2	8.6	2.0
3a	8.0	8.8	8.4	9.3	12.6	16.6	18.7	19.0	15.5	9.5	6.7	2.4
3b	5.8	7.6	8.6	9.1	10.0	9.7	9.3	8.3	7.4	6.7	8.3	2.0
4a S	6.8	6.5	5.9	6.7	6.9	6.8	7.0	8.2	9.6	8.2	9.9	1.7
4a P	7.8	5.2	4.9	4.4	3.9	4.3	4.1	3.9	7.8	6.6	7.1	1.9
4b	12.0	9.1	6.7	4.5	4.2	3.2	3.6	4.7	7.2	8.6	9.6	2.2

Table 2. Wuyi Citation Tones: Standard Deviation Figures

Table 2 presents standard deviation figures which correspond to the mean values presented in Table 1. All standard deviation figures are given to one decimal place. Thus, for example, we can see that Wuyi citation tone 1b has a mean value of 124 Hz at the 70% sample point with a corresponding standard deviation value of 5.9 Hz. Tone 1b's mean duration is 28.9 csecs with a corresponding standard deviation value of 1.9 csecs.

DISCUSSION

Wuyi Unstopped ('Shu') Citation Tones

Tone 1a, graphed in Figure 2 as a function of time, has low rising pitch. Graphically, we see the mean F_0 contour of Tone 1a in citation form is slightly concave, dipping a little before beginning to rise at around 30% of its mean duration. The mean F_0 contour begins at a value of 122 Hz, dips slightly by about 8 Hz to 114 Hz, then reaches a maximum mean frequency of 160 Hz at offset. The mean duration is 25.0 csecs.

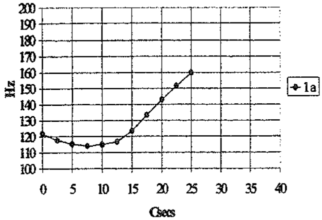

Figure 2. Wuyi Citation Tone 1a

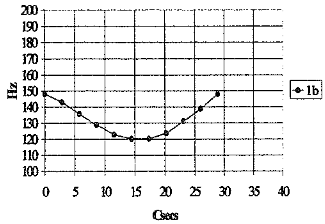

Figure 3. Wuyi Citation Tone 1b

Tone 1b, shown in Figure 3, has mid fall-rising pitch in citation form, which is clearly visible here in its graphical representation. The mean citation contour begins at a value of 148 Hz, falling to its lowest value of 120 Hz between 50% and 60% duration. The contour then returns to a maximum mean frequency of 148 Hz at offset. The mean duration of Tone 1b's citation allotone is 28.9 csecs, rendering it on average, the longest citation contour of Wuyi.

Tone 2a, Figure 4, has high level-rising pitch in citation form. Graphically, we see the mean underlying F_0 contour remains roughly level at approximately 155 Hz for the first 40%-50% of its mean duration. After this point, F_0 begins to rise somewhat linearly for the remainder of the duration, reaching a mean value of 168 Hz by offset. Mean tonal duration is 22.2 csecs. There is no reason apparent in the segment structure to assume that the rising portion of this contour is not extrinsic.

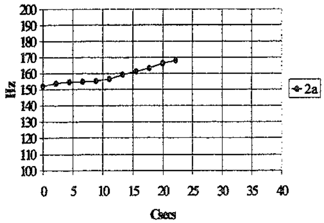

Figure 4. Wuyi Citation Tone 2a

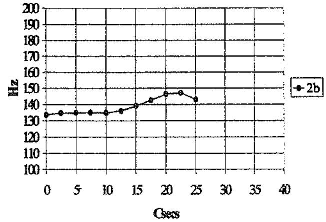

Figure 5. Wuyi Citation Tone 2b

Tone 2b, Figure 5, has mid level-rising pitch in citation form. Graphically we can see that the mean F_0 contour of Tone 2b in citation form is very similar in shape to the other level-rising pitched tone that we have just seen, however, realised with lower F_0. The mean F_0 contour remains roughly level at approximately 135 Hz for the first 40%-50% of the contour's mean duration. After that point, F_0 begins to rise somewhat linearly with similar slope to that exhibited by Tone 2a and reaches a peak mean value of 147 Hz. Unlike Tone 2a, all tokens of Tone 2b consistently demonstrate a tendency towards a small drop in F_0 in the last 10%-20% of their duration. This drop amounts to a mean fall of just 4 Hz, prior to the offset of phonation. Tone 2b has a mean duration of 25.1 csecs. Again, there is no reason to assume that the rising portion of this tone is an intrinsic feature.

Tone 3a, Figure 6, has high falling pitch. Graphically, we can see that the mean F_0 contour of Tone 3a in citation form is slightly convex or sigmoid in shape. Beginning at a mean value of 177 Hz, mean F_0 rises very slightly by 1 Hz over the first 10%, remains level at 178 Hz over the next 10% then begins to fall sharply with acceleration until reaching about 50% of total mean duration. The contour then falls more or less with deceleration during the latter 50%, reaching a minimum value of 126 Hz at offset. The mean duration is 16.1 csecs. Many of the individual tokens of this contour showed a rise in F_0 during the last 10%-30% of duration prior to offset, others were simple falls. This rise was, however, not audible as pitch.

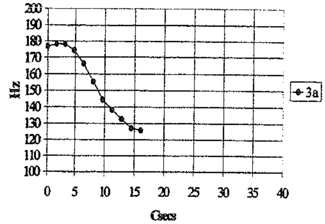

Figure 6. Wuyi Citation Tone 3a

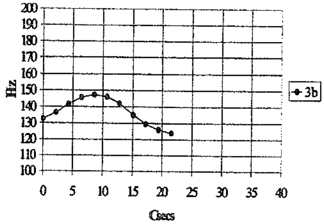

Figure 7. Wuyi Citation Tone 3b

Tone 3b, Figure 7, has mid rise-falling pitch in citation form. The mean F_0 contour begins with a fundamental frequency of 133 Hz, climbs to a peak value of 147 Hz at 40% of total mean duration, then falls to a value of 124 Hz at phonation offset. Mean duration is 21.5 csecs.

Wuyi Stopped ('Ru') Citation Tones

Morphemes which historically carried Tones belonging to the Ru or stopped tone categories 4a and 4b have undergone or are undergoing various changes in Wuzhou. Many such morphemes have lost their stopped codas, either in citation form or in Sandhi form or both. Some of the tonal contours realised on these morphemes have merged with one or another of the tonemes within the dialect. Thus, when eliciting morphemes on the basis of their belonging to this historical category, we are often presented with a number of different syllable types and tonal shapes. We find this situation to be true of Wuyi.

Examples belonging to the historical category 4a in Wuyi exhibit two tonetically distinct tonal shapes shown in Figure 8 and Figure 9. Those syllables with clearly audible glottal stops in coda position appear to exhibit one shape, Figure 8, and those with less clearly audible or inaudible glottals appear to follow another shape, Figure 9, which is merging or has merged with citation tone 3a. Both contours have high falling pitch tonologically, however, those from the group with strong glottal stops, which I refer to as contour 4a S (Short), have a tonetically truncated shape which never reaches as low an F_0 value by phonation offset as do examples of the group with less clearly realised glottals, which I refer to as contour 4a P (Protracted).

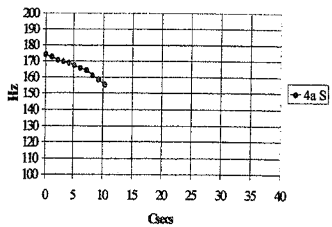

Figure 8. Wuyi Citation Tone 4a S

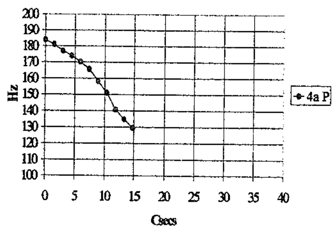

Figure 9. Wuyi Citation Tone 4a P

Tone 4a S, Figure 9, has truncated high falling pitch. Graphically, we see that it begins at a mean value of 174 Hz and then falls sharply to 156 Hz at offset. The mean duration is 10.3 csecs. The falling pitch contour realised upon this syllable type is a little shallower than that realised on syllables carrying 4a P. We could attribute the tonetic differences we observe between this contour and that of 4a P to features of the truncation undergone due to the reduced nucleus length, which itself is an intrinsic effect brought about by the closing of the syllable by an obstruent. The most salient auditory feature of this contour, rendering it identifiable from that of contour 4a P is its duration rather that any F_0 differences.

Tone 4a P, Figure 9, has high falling pitch. Its F_0 contour begins at a mean value of 184 Hz, then F_0 falls sharply to 130 Hz at offset: The mean duration is 14.8 csecs. This contour is auditorily indistinguishable from the high falling contour exhibited by Tone 3a. Statistical analysis of these contours by ANOVA shows that any observed difference between the contours is too likely due to chance to reject the null hypothesis and thus we must assume homogeneity. However, it is interesting nonetheless to note that visual inspection of individual contours suggests some consistent trends (notably the more linear sloping onset of 4a P versus the more convex onset of 3a) which would be difficult to attribute to the intrinsic effects of segment structure or the like and which the author speculates may be the result of the ongoing lengthening of 4a S and conjectures, given more data, might be demonstrable as indicative of the separate origins of the two contours. No further reference to this contour will be made in this paper, as it is irrelevant to the paper's central argument.

Tone 4b, Figure 10, has truncated mid fall-rising pitch in citation form. The mean F_0 contour begins at a value of 143 Hz. Falling to its lowest value of 122 Hz at the 50% sample point, it climbs to 133 Hz and remains at this value between the 80% and 90% sample points before shelving off slightly to 132 Hz at offset. The mean duration of Tone 4b's citation allotone is 25.0 csecs. We might consider the underlying tone command behind Tone 4b's citation allotone to be the same as that of Tone 1b under the influence of intrinsic tonetic modifications brought about by the glottal closure of the syllable.

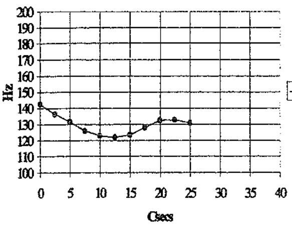

Figure 10. Wuyi Citation Tone 4b

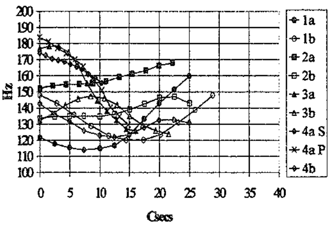

Figure 11. Wuyi Citation Tones

Tonological Implications

The mean acoustic data presented in Table 1 clearly confirms that the three rises observed auditorily in the Wuyi citation tone system are indeed present. Figure 11 shows all Wuyi citation tones graphed together as functions of time for comparison. Graphically, Tone 1a, is perhaps more typical amongst low rising tones, having a shallow fall at onset which is inaudible as pitch, followed by a salient rise component and an overall concavity in appearance. Tones 2a and 2b, however, exhibit F_0 contours which remain largely level over the first half of their duration but which rise over the latter half. In both cases, the level component as well as the rising component are audible as pitch. These rise components represent mean rises of 13 Hz and 12 Hz or 18.6% and 17.1% of the speaker's pitch range respectively (as defined in terms of the data presented here) and cannot be attributed to the effects of segment structure. This finding poses serious questions for modern tonological models such as Yip (1980) and her subsequent revisal, Yip (1989), which are based on the tenet that only two 'complex tones' (dynamic tones) of any one contour type may appear in contrastive distribution.

TONE	TONE CATEGORY	CONTOUR	TONE LETTERS	REGISTER TIER	MELODIC TIER
1a	Yinping	low rising	[14]	-U	LH
1b	Yangping	mid fall-rising	[313]	-U	HLH
2a	Yinshang	high level-rising	[334]	+U	LH
2b	Yangshang	mid level-rising	[223]	-U	LH??
3a	Yinqu	high falling	[51]	+U	HL
3b	Yangqu	mid rise-falling	[231]	-U	LHL
4a	Yinru	truncated high falling	[53ʔ]	+U	HL
4b	Yangru	truncated mid fall-rising	[312ʔ]	-U	HLH

Table 3. Underlying Representations under Yip (1980, 1989)

Table 3 summarises the Wuyi citation tones, presenting pitch descriptors, column 3, Chao's (1930) tone letters, column 4, and proposed underlying representations under the Yip model, columns 5 and 6. Table 3 also highlights the representational problems Wuyi poses for the Yip model. While the ru contours might be considered allotonic realisations of other tonemes conditioned by glottal finals and thus, may have the same underlying representations, Tone 2b cannot be kept distinct from Tone 1a which occurs in similar environments. Under such a model, it is not possible to keep three rising contours distinct without either recourse to the branching root node structures proposed in Yip (1989) or without relying on the ability to abstract by positing underlying forms whose structures fill unused slots within the matrix of the model and whose derivational forms cannot be expressed in terms of that model. The branching root node structures argued for in Yip (1989) which the author does not support and has presented evidence against in Thompson (1998) are presented primarily as a tool by which to explain the simplification of tone structures under sandhi but may also serve to circumvent the OCP of Leben (1973) and Goldsmith (1976), effectively allowing sequences of tones by employing a structure such as that shown in Figure 12. These tonological issues are explored further in Thompson (in preparation).

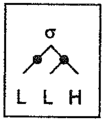

Figure 12.

CONCLUSION

To summarise, this paper shows that Wuyi contrasts 8 contours in citation form, 3 rising, 2 falling, 2 fall-rising and 1 rise-falling. The results thus, confirm the initial supposition of this paper that there exist three distinct rising F_0 shapes in contrastive distribution on citation syllables within the Wuyi dialect.

ACKNOWLEDGEMENTS

I should like to thank Prof. William L. Ballard whose data recorded in the field served as the raw material for this analysis and Dr. Philip J. Rose and my reviewers for their comments.

REFERENCES

Chao, Y-R. (1930) *A System of Tone Letters*, Le Maître Phonétique 45: 24-47.

Goldsmith, J. (1976) *Autosegmental Phonology*, Ph.D. dissertation, MIT, Published 1979, (Garland: New York).

Leben, W. R. (1973) *Suprasegmental Phonology*, Ph.D. dissertation, MIT, (Distributed by Indiana University Linguistics Club: Bloomington, Indiana).

Rose, P. J. (1990) *Acoustics and Phonology of Complex Tone Sandhi*, Phonetica 47: 1-35.

Rose, P. (1998) *The Differential Status of Semivowels in the Acoustic Phonetic Realisation of Tone*, In Robert H. Mannell & Jordi Robert-Ribes (eds.) Proceedings of The 5th International Conference on Spoken Language Processing (CD Rom), (Causal Productions: South Australia).

Thompson, N. G. I. (in preparation) *(Yet to be Titled)*, Ph.D. Thesis, Australian National University, Canberra.

Thompson, N. G. I. (1998) *Tone Sandhi between Complex Tones in a Seven-Tone Southern Thai Dialect*, In Robert H. Mannell & Jordi Robert-Ribes (eds.) Proceedings of The 5th International Conference on Spoken Language Processing (CD Rom), (Causal Productions: South Australia).

Yip, M. J. W. (1980) *The Tonal Phonology of Chinese*, Ph.D. dissertation, MIT, (Distributed by Indiana University Linguistics Club: Bloomington, Indiana).

Yip, M. J. W. (1989) *Contour Tones*, Phonology 6, pp. 149-74, (C.U.P.).

Zhu, S. & Rose, P. (1998) *Tonal Complexity as a Dialectal Feature: 25 Different Citation Tones from Four Zhejiang Wu Dialects*, In Robert H. Mannell & Jordi Robert-Ribes (eds.) Proceedings of The 5th International Conference on Spoken Language Processing (CD Rom), (Causal Productions: South Australia).